An Investigation of Institutional Characteristics Associated with Response Rates in Mail Surveys of Community Hospitals

ROSS MULLNER, PhD
PAUL S. LEVY, ScD
DALE MATTHEWS, BS
CALVIN S. BYRE, PhD

COMMUNITY HOSPITALS, because of their great importance as providers of inpatient and outpatient medical care in the United States, are frequently the subject of studies conducted by a variety of health professionals. Generally these researchers seek information by means of survey questionnaires mailed to the community hospitals (defined by the American Hospital Association as all non-Federal short-term general and other special hospitals—excluding hospital units of institutions—whose facilities and services are open to the public). As with all surveys, the validity of these workers' findings is related to the rates of response: findings from surveys that have poor rates of response must be considered suspect.

Although a large number of such mail surveys have been conducted, very little is known about the factors associated with hospitals' rates of response. Clearly, there are a variety of factors that influence or determine response rates: for example, the reputation and professional standing of the sponsor or conductor of the survey, the conception and design (length, number of questions, and so forth) of the survey instrument, and characteristics of the recipient. Most of these factors do not readily lend themselves to statistical analysis or permit general, verifiable conclusions, for they vary from survey to survey and are difficult to define and to quantify objectively.

Many characteristics of recipients, however, are relatively constant and are objectively definable and quantifiable. It seems reasonable, therefore, for research into factors associated with rates of response to surveys to begin with these.

It is the purpose of this study (*a*) to describe the nature and strength of the relationship between certain institutional characteristics of the U.S. community hospitals listed with the American Hospital Association and their rates of response to a nationwide AHA survey and (*b*) to propose a method of predicting, on the basis of these characteristics, whether a hospital is likely to respond to surveys. Because it is believed that the patterns of response to AHA surveys are paralleled in surveys of hospitals conducted by other organizations, these findings should be of broad interest, and the method of predicting response should be useful to other researchers with similar concerns, enabling them to improve overall rates of response by concentrating their more intense (and generally more expensive) efforts on those hospitals identified as likely nonresponders.

Methods

Data source. The data source for this study is the 1978 Special Survey on Selected Hospital Topics con-

Dr. Mullner, Mr. Matthews, and Dr. Byre are with the Hospital Data Center of the American Hospital Association. Dr. Levy is with the University of Illinois School of Public Health at the Medical Center, Chicago. Tearsheet requests to Ross Mullner, PhD, Hospital Data Center, American Hospital Association, 840 N. Lake Shore Dr., Chicago, Ill. 60611.

ducted by the AHA. This annual survey consists of a questionnaire mailed to the chief executive officer of every U.S. community hospital that is listed in the AHA's computer file of hospitals (the file is updated weekly and is considered to be the most complete and up-to-date list available). The single-page questionnaire contains 10 to 15 questions; the subject matter varies from year to year, but the size of the questionnaire remains the same. The 1978 survey sought information about hospital liability insurance, audiovisual equipment, size of the medical staff, revenue and deductions, equipment expenses, and so forth.

Variables examined. From the computer tape used to process the data from the 1978 survey, we prepared a working tape containing information about each of the 5,867 community hospitals to which a questionnaire was mailed. This information consisted of six variables chosen because they are the ones most widely used by health organizations in characterizing and classifying hospitals:

1. Number of beds
2. Whether or not the hospital is a teaching hospital
3. Whether or not the hospital is an AHA member
4. Whether or not the hospital is located in a Standard Metropolitan Statistical Area (SMSA)
5. Whether or not the hospital is investor-owned
6. Whether or not the hospital is controlled by a State or local government

Also recorded on the working tape was whether or not the hospital responded to the survey questionnaire.

Statistical methods and analysis. In the first stage of our analysis we examined separately the effects of each of the six characteristics on response. Of the six, all are dichotomies except number of beds; this variable was made dichotomous for this stage of the analysis by stratification of the hospitals into two classes: those with fewer than 200 beds and those with 200 or more. We next constructed fourfold tables that cross-tabulated response status by each of the six dichotomous variables. From each fourfold table we then computed the standard chi-squared statistic (denoted X^2) which tests the null hypothesis that the response rate is the same for each level of the variable. In addition, the strength of the association of response with the variable was measured by the odds ratio for each fourfold table. The odds ratio, a commonly used measure of association (*1*), is computed by taking the ratio of the cross-products of the entries in the cells of the fourfold table. In this instance, it represents the estimated odds of a hospital responding to the questionnaire when a certain characteristic (for example, AHA membership) is present as opposed to when it is absent.

Table 1 shows the rates of response of community

Table 1. Number of hospitals, response rates, significance of X^2 statistic, and response odds ratios overall and stratified according to hospital characteristics: AHA Special Survey on Selected Hospital Topics, 1978

Stratification	Number of hospitals	Response rate (percent)	Significance of X^2 statistic (P value)	Response odds ratio (calculated from fourfold tables)
Overall	5,867	69.4	...	...
Number of beds:				
Fewer than 200	4,217	64.0	<.001	2.79
200 or more	1,650	83.3		
Teaching status:				
Teaching	947	84.6	<.001	2.76
Nonteaching	4,920	66.5		
AHA membership:				
Members	4,652	75.7	<.001	3.79
Nonmembers	1,215	45.2		
Location of hospital:				
Standard Metropolitan Statistical Area	2,919	73.5	<.001	1.46
Non-Standard Metropolitan Statistical Area	2,948	65.4		
Ownership category:				
Investor owned	742	43.9	<.001	[1]0.24
State and local government	1,787	66.6	<.001	[1]0.59
Nongovernment, not-for-profit	3,338	76.8	<.001	[2]4.23

[1] Odds ratios are relative to nongovernment not-for-profit hospitals.

[2] Odds ratio is relative to investor-owned hospitals.

hospitals to the 1978 AHA special topics survey, both overall and grouped according to the characteristics examined in this study. Overall, 69.4 percent of the 5,867 hospitals responded to the survey. Response rates were significantly higher among hospitals having 200 or more beds (83.3 percent) than among hospitals having less than 200 beds (64.0 percent); among teaching hospitals (84.6 percent) than among nonteaching hospitals (66.5 percent); among members of the AHA (75.7 percent) than among nonmembers (45.2 percent); and among hospitals located in SMSAs (73.5 percent) than among those located in nonmetropolitan areas (65.4 percent). Of the private, not-for-profit hospitals, 76.8 percent responded as opposed to 66.6 percent of the hospitals controlled by State and local governments and 43.9 percent of the investor-owned hospitals.

Table 1 also shows the response odds ratios calculated from the fourfold tables. Indicating the strength of the apparent association between an independent variable and response, the ratios show that, for example, the odds of a hospital with 200 or more beds responding to the survey is 2.79 times that of a hospital with fewer than 200 beds. The highest values of these odds ratios are for AHA membership (3.79) and for nongovernment not-for-profit control (4.23).

These odds ratios are, however, subject to confounding biases. Confounding biases can occur when two variables, A and B, are each associated not only with a third, dependent variable, C, but also with each other, and when only one of these, A, is a determinant of that third variable. In such a case, the method just described might lead to the erroneous conclusion that B is a determinant of C.

Of the wide variety of methods that can be used to detect and control the effects of confounding, we used multiple logistic regression; it is particularly useful when the dependent or outcome variable is dichotomous (such as response versus nonresponse to a survey) and the set of independent variables is a mixture of categorical and quantitative members. In this method, which has been widely used by epidemiologists in etiological studies identifying the risk factors that increase the likelihood of developing chronic diseases (2), the relationship between a dichotomous dependent variable, Y, and a set of independent variables, $X_1, X_2, \ldots, X_k$ is specified by the equation:

$$Y = e^{(b0 + b_1X_1 + \ldots + bk\,Xk)} \div (1 + e^{(b0 + b_1X_1 + \ldots bk\,Xk)})$$

Each regression coefficient, b_j, in this model measures the strength of the relationship between the independent variable X_j, and the dichotomous dependent variable Y, while controlling for possible confounding due to the other independent variables.

From the values of the set of independent variables substituted into the equation, a number between zero and unity is computed; this number is an estimate of the probability that the dichotomous dependent variable takes a specified value. In our application, it estimates

the probability that a hospital having characteristics specified by the set of independent variables will respond to the survey.

Using a computer program based on the method of Walker and Duncan (*3*), we fitted our variables to the multiple logistic equation, specifying these variables as follows:

Variable	*Coefficient*	*Specification*
Constant	b_0	
Number of beds (X_1)	b_1	Actual number of beds
Teaching status (X_2)	b_2	1 if teaching hospital; 0 if nonteaching hospital
AHA membership (X_3)	b_3	1 if AHA member; 0 otherwise
Location of hospital (X_4)	b_4	1 if located in SMSA; 0 otherwise
Government controlled (X_5)	b_5	1 if controlled by a State or local government; 0 otherwise
Investor-owned (X_6)	b_6	1 if investor-owned; 0 otherwise

Number of beds is the only quantitative variable in the equation. The other variables are categorical; of these, AHA membership, hospital location, and teaching status are already dichotomous, whereas the three categories of hospital ownership have been recoded into two dichotomous variables (the usual procedure in multiple regression analysis for categorical variables having more than two categories).

Table 2 shows the magnitude of the estimated regression coefficients relative to their standard errors, and the statistical significance of the relationship between response and each independent variable when there is control for possible confounding due to the other variables.

Table 2 also shows the logistic odds ratio for each independent variable derived from the equation; these ratios are estimates of the independent effect that a specified value of a variable has on the odds of responding. They can be compared to the "crude" response odds ratios transcribed from table 1, where there was no control for possible confounding effects from the other variables. Of the variables that appeared from the fourfold table analysis to be significantly related to response, all except teaching status had significant regression coefficients. It appears, then, that the relationship shown in table 1 between teaching status and response was a spurious one caused by confounding with the other variables. However, the relationships indicated in table 1 between response and the remaining variables—AHA membership, type of ownership, and bed size of the hospital—all remain strongly related to response when adjustment is made for confounding. Thus, for example, the odds of an AHA member hospital responding to the survey is 2.49 times that of an AHA nonmember, ceteris paribus.

The ability of the logistic model to predict response is shown in table 3. The estimated probability of response of each of the 5,867 community hospitals in the 1978 AHA special topics survey was computed from the equation based on the coefficients shown in table 2. The hospitals were then grouped into eight classes according to probability of response, and the actual response rates were calculated for each class. As shown in table 3, actual response rates increase at an equal pace with increases in the estimated logistic response probabilities ($r = .98$, $P < .01$ over the 8 classes). For example, only 25 percent of the 40 hospitals having logistic response probabilities between 0.20 and 0.29 actually responded, whereas 83.2 percent of the 149 hospitals having estimated logistic response probabilities between 0.90 and 0.99 responded.

Table 4 shows the results of the same procedure applied to two other randomly selected surveys of commu-

Table 2. Estimated multiple logistic regression coefficients and response odds ratios as estimated from fourfold tables and from the multiple logistic: AHA Special Survey on Selected Hospital Topics, 1978

Variable	*Regression coefficient*	*Significance of regression coefficient (P value)*	*Estimated response odds ratios*	
			Calculated from fourfold table	*Estimated from multiple logistic*
Constant	$b_0 = -.022683$	NS	...	...
Number of beds	$b_1 = .0017552$	<.001	[1]2.79	[1]1.56
Teaching status	$b_2 = .098783$	NS	2.76	1.10
AHA membership	$b_3 = .91367$	<.001	3.79	2.49
Location of hospital	$b_4 = .16556$	<.05	1.46	1.18
Government controlled	$b_5 = -.23258$	<.001	[2]0.59	[2]0.79
Investor owned	$b_6 = -.88517$	<.001	[1]0.24	[2]0.41

[1] 200 or more beds versus fewer than 200 beds.
[2] Odds ratios are relative to nongovernment not-for-profit hospitals.

Note: NS = not significant.

Table 3. Total number of hospitals and percentage responding grouped according to probability of responding as estimated from the logistic: AHA Special Survey on Selected Hospital Topics, 1978

Probability of responding as estimated from the logistic	Total number of hospitals	Percent responding
0.20–0.29	40	25.0
0.30–0.39	356	38.8
0.40–0.49	424	44.1
0.50–0.59	516	50.2
0.60–0.69	918	64.6
0.70–0.79	2,240	77.0
0.80–0.89	1,224	84.8
0.90–0.99	149	83.2
Total	5,867	69.4

nity hospitals, surveys that differ in certain important respects, such as size, source, and subject, from the special topics survey. The Survey of Hospital Governing Boards of March 1979 was conducted by the AHA; the 8-page questionnaire, mailed to 5,815 hospitals, asked for information about the bylaws and other requirements regulating their governing boards, the boards' composition and members (demographic characteristics, financial compensation, and so forth). The 1979 Survey of Hospital Wages was conducted by Vanderbilt University; the 4-page questionnaire was mailed to a representative sample of 1,601 hospitals and asked about the wage scale for each type of position among hospital employees.

As in table 3, hospitals were grouped into eight classes according to the logistic probability of response, and the actual response rates are shown for each class. The table supports the findings about the response rates of the special topics survey described previously. Again (except where the responses for a class were too few to yield a stable rate) the actual response rates increase with increases in the probability of response. The correlations between the actual and estimated responses were all highly significant ($P < .01$ over the 8 classes). The correlation was $r = .94$ for the Survey of Hospital Governing Boards and $r = .95$ for the Survey of Hospital Wages.

Of the five characteristics that have a significant relationship to response, three—AHA membership, bed size of 200 or more, and location in an SMSA—are correlated with relatively high rates of response. The other two—investor ownership and government control—are correlated with relatively low response rates. Why these relationships exist is difficult to determine; the following are possible explanations.

Variables Associated with Low Response Rates

Investor ownership. Many investor-owned hospitals are owned or managed by corporations and must redirect requests for information to their central offices; this process sometimes takes a long time, and the survey can be misdirected or ignored. Furthermore, corporate offices seems generally to be less inclined to release information and to accept the cost of participating in survey projects than not-for-profit institutions.

State and local government control. It is often the case that some of the records of these hospitals are kept at some distance from them in central government offices; thus, as is the case with investor-owned hospitals, the process of collecting information becomes complicated and difficult.

Variables Associated with High Response Rates

AHA membership. Staffs of AHA member hospitals have a commitment to the goals of the Association, in-

Table 4. Total number of hospitals and response rates to two surveys grouped according to probability of responding as estimated from the logistic

Probability of responding as estimated from the logistic	AHA Survey of Hospital Governing Boards		Vanderbilt University Survey of Hospital Wages	
	Total number of hospitals	Percent responding	Total number of hospitals	Percent responding
0.20–0.29	39	33.3	0	0.0
0.30–0.39	362	44.2	69	20.3
0.40–0.49	444	61.9	21	33.3
0.50–0.59	507	58.6	63	28.6
0.60–0.69	902	76.1	57	38.6
0.70–0.79	2,189	82.4	494	42.7
0.80–0.89	1,218	86.3	771	54.3
0.90–0.99	154	81.2	126	69.0
Total	5,815	75.9	1,601	48.6

cluding that of collecting and disseminating health care information; furthermore, since this information is regularly made available to them, they have a direct interest in facilitating the process of gathering it.

Bed size of 200 or more. These hospitals, unlike the smaller ones, are more likely to have computer systems that make it relatively easy to store and retrieve information about their operations. Many of them have established procedures for handling surveys quickly and efficiently, and some even have staff members whose sole function is to deal with them.

Location in SMSA. Perhaps the hospitals located in SMSAs respond at relatively higher rates than non-SMSA hospitals because metropolitan hospital associations actively encourage participation in surveys. Or perhaps because these hospitals are generally geographically closer to, and their officers have more frequent contacts with, professional organizations and political bodies, the officers believe strongly that the information gathered by surveys is used in making policies or in proposing legislation that directly affects their institutions.

Conclusion

This study has shown that five primary characteristics of community hospitals—bed size, SMSA-non-SMSA location, AHA membership, ownership, and control—are significantly related to hospitals' response rates to surveys. Moreover, it has shown how the relative probabilities of response for hospitals possessing various combinations of these characteristics can be accurately predicted.

The methods we have described appear to be of general applicability, and they could easily be used to investigate the relationship to survey response of many other features or characteristics shared by groups of hospitals.

These methods are also immediately useful. By employing them, researchers can predict more accurately which groups of hospitals are likely to contain a large proportion of nonresponders to their surveys and can therefore seek to improve response rates by concentrating special efforts on these groups. For example, before sending a survey questionnaire, researchers could call or even visit these hospitals to explain its purpose and show why it is important. They may enlist the aid of local hospital associations or of corporations controlling investor-owned hospitals to persuade the staffs of these hospitals to participate in the project. Or they may direct a greater number of followup mailings to the group of potential nonresponders.

References

1. Fleiss, J. L.: Statistical methods for rates and proportions. John Wiley & Sons, Inc., New York, 1973.
2. Prentice, R. L.: Use of the logistic model in retrospective studies. Biometrics 32: 599–606 (1976).
3. Walker, S. H., and Duncan, D. B.: Estimation of the probability of an event as a function of several independent variables. Biometrika 54: 167–179 (1967).

SYNOPSIS

MULLNER, ROSS (American Hospital Association), LEVY, PAUL S., MATTHEWS, DALE, and BYRE, CALVIN S.: *An investigation of institutional characteristics associated with response rates in mail surveys of community hospitals. Public Health Reports, Vol. 96,* March–April 1981, pp. 128–133.

This paper describes the nature and strength of the relationship between six institutional characteristics of U.S. community hospitals and the rates of response of these hospitals to a nationwide survey conducted by the American Hospital Association (AHA). Furthermore, it demonstrates how one can calculate accurately the relative probability of response of hospitals with various combinations of these characteristics. The six characteristics studied were bed size, teaching status, AHA membership status, location within or without a Standard Metropolitan Statistical Area, investor or other form of ownership, and control by State or local government or by another type of organization.

The six characteristics were treated as dichotomous variables throughout most of the analysis. Odds ratios were calculated for each variable as a preliminary measure of the strength of its association with response.

The effects of confounding on those odds ratios were controlled for by multiple logistic regression, which estimates the probability of response of hospitals with given characteristics. A logistic odds ratio was calculated for each variable to estimate the independent effect that specified values had on the odds of responding. All variables except status as a teaching or nonteaching hospital were shown to have a significant relationship to response.

Hospitals were divided into classes according to probability of response, and the probability of response was compared with actual rates of response in the survey. Both increase at equal rates. Similar results were obtained by applying this method to two other, dissimilar, surveys.

This method is generalizable to other surveys and should prove useful to researchers wishing to increase response rates in their surveys of hospitals.